

Singes et hommes face au monde visuel : la catégorisation

Michèle Fabre-Thorpe, Arnaud Delorme
et Ghislaine Richard

Centre de Recherche Cerveau et Cognition, UMR 5549, Toulouse, France

Résumé

On sait peu de choses sur les mécanismes cérébraux sous-jacents aux capacités de catégorisation de l'animal. Les pigeons pourraient baser leurs réponses sur des caractéristiques simples des stimulus, mais qu'en est-il des singes ? Nous avons comparé les capacités de catégorisation rapide de 2 groupes de 10 hommes et de 2 macaques rhésus dans une tâche où les sujets doivent détecter la présence d'une cible donnée (soit animal, soit aliment) dans une image naturelle qui n'est présentée que pendant 30-80 ms. Hommes et singes sont testés sur les mêmes ensembles d'images. Les hommes travaillent sur des séries de 100 images nouvelles. Les macaques sont testés chaque jour sur 5 nouvelles cibles et 5 nouveaux distracteurs présentés parmi des images familières. Seule la réponse donnée par le singe à la *première* présentation d'une image nouvelle est analysée. Sur différents ensembles de 200 images nouvelles les résultats sont reproductibles, les macaques se révèlent légèrement moins précis (de 84% à 90.5% de réussite) que les hommes (95% de réussite en moyenne), mais beaucoup plus rapides (TR médian variant de 233 à 344 ms chez le singe et de 383 à 504 ms pour les hommes). Une analyse spécifique menée sur les erreurs a montré que hommes et singes manifestaient des difficultés face aux mêmes images. De plus, apprentissage et manipulation des indices de couleur affectent leurs performances de façon similaire. La familiarisation avec les images augmente la précision sans raccourcir la latence des réponses les plus précoces. L'absence d'indices de couleur ne perturbe ni les performances des macaques ni

Adresse de correspondance : Centre de Recherche Cerveau et Cognition (UMR 5549) CNRS, UPS, Faculté de Médecine de Rangueil, 133 route de Narbonne, 31062 Toulouse, France (e-mail: mft@cerco.ups-tlse.fr).

celles des hommes produisant de rapides réponses (TR courts). En revanche, la couleur semble être prise en compte dans les décisions des hommes produisant des réponses lentes. Hommes et singes pourraient être dotés de processus de catégorisation ultrarapides remarquablement efficaces. Les mécanismes cérébraux sous-jacents à ce type de catégorisations rapides pourraient présenter de grandes similarités chez l'homme et le singe et être basés sur le traitement unidirectionnel des premières informations visuelles disponibles.

Mots clés : Macaques, hommes, scènes naturelles, catégorisation visuelle, traitement visuel.

Key words: Macaques, humans, natural scenes, visual categorisation, visual processing.

INTRODUCTION

L'homme organise tout naturellement ses connaissances en catégories et regroupe sous un même nom des idées, des événements ou des objets très dissemblables. Il n'y a pas si longtemps, certains auteurs affirmaient d'ailleurs que seul l'homme était doté de telles capacités de catégorisation : "Unlike any other animal, we have a natural ability to group objects or events into categories" (Hunt, 1982, cité dans Herrnstein, 1984). D'autres admettaient que certains animaux en soient pourvus, mais la possibilité de faire appel à des représentations abstraites de catégories d'objets n'était accordée qu'aux anthropoïdes ayant eu accès au langage (Premack, 1983). Pourtant, identifier un nouvel objet comme appartenant à une catégorie dont on connaît les propriétés (prédateur, proie, aliment, congénère, ennemi ...) est un avantage biologique incontestable. Catégoriser est un moyen simple pour traiter et organiser les millions d'informations qu'un animal reçoit chaque jour !

C'est le travail original de Herrnstein et Loveland, en 1964, qui a véritablement ouvert la voie à l'étude des capacités de catégorisation chez l'animal. Ce travail présentait d'étonnants résultats obtenus chez le pigeon, qui s'est révélé non seulement capable de reconnaître dans un lot de diapositives celles où figurait un être humain, mais qui a également présenté la capacité de transférer cet apprentissage sur un nouvel ensemble de diapositives jamais vues auparavant. Depuis, les étonnantes facultés de classification de ces oiseaux ont été mises en évidence pour les arbres, les feuilles des différentes espèces de chênes, les poissons ou même des objets arbitrai-

res comme des A ou des 2. Ce travail et les études qui ont suivi ont donné naissance à une controverse toujours très vivace sur la nature des concepts qu'un animal peut acquérir. Il est surprenant de voir qu'il a fallu attendre les années 80 pour que des études similaires soient menées chez le singe. Ces expériences montrent que les singes sont capables d'apprendre à catégoriser un large éventail de stimulus (D'Amato et Van Sant, 1988 ; Roberts et Mazmanian, 1988 ; Schrier, Angarella et Povar, 1984 ; Schrier et Brady, 1987 ; Yoshikubo, 1985). Elles montrent également que ces capacités de catégorisation peuvent être utilisées dans des conditions naturelles. Ainsi le singe vervet émet un cri d'alarme différent en fonction du prédateur repéré (léopard, aigle ou python), et provoque une réaction de fuite adaptée chez ses congénères (Seyfarth, 1980). Mais les mécanismes cérébraux mis en jeu pour effectuer ces catégorisations restent mal connus. Un pigeon catégorise-t-il de la même façon qu'un singe ? Qu'en est-il de l'homme ? Si la capacité de catégoriser n'est apparue qu'une seule fois au cours de l'évolution, elle nous viendrait d'un ancêtre commun très ancien puisque mammifères et oiseaux en sont pourvus. En revanche, elle a pu apparaître plusieurs fois et de façon indépendante dans différents groupes animaux. Dans ce cas, le type de catégorisation que nous partageons avec les singes viendrait d'un ancêtre commun "récent" et les processus de catégorisation mis en jeu chez le singe pourraient être tout à la fois différents de ceux utilisés par le pigeon et similaires – au moins pour les processus de base – à ceux de l'homme ! De récentes études menées par l'équipe de Huber (Aust, Loidolt et Huber, 1999 ; Troje, Huber, Loidolt, Aust et Fieder, 1999) ont montré que le pigeon semble baser ses réponses sur des caractéristiques simples des stimulus, la couleur, la texture et la luminosité jouant un rôle important alors que la forme semble ignorée. Mais en ce qui concerne les singes, peu de travaux se sont attachés à déterminer les similitudes et les différences des mécanismes sous-tendant les catégorisations qu'ils effectuent avec celles effectuées par l'homme. Bien que certaines études comparatives aient été menées dans le domaine de l'audition (Owren, 1990a, 1990b), c'est à la modalité visuelle que seront ici empruntés les exemples en s'attachant plus précisément à la comparaison des performances de l'homme et du singe.

DE LA DISCRIMINATION A LA CATÉGORISATION

Il ne peut être ici question d'être exhaustif, mais plus modestement de montrer que certains singes – et tout particulièrement des singes de l'ancien monde, comme le macaque rhésus – possèdent les aptitudes de base indis-

pensables pour pouvoir développer des "représentations mentales" de différentes classes d'objets et, ceci, sans préjuger de la nature de ces représentations.


L'une des premières questions concerne leur capacité à apprendre une règle générale. Placés face à divers problèmes solubles par l'application d'une même stratégie, les singes sont-ils capables d'apprendre cette stratégie ou sont-ils obligés de résoudre chaque problème en associant une réponse donnée à un stimulus précis ? Si l'on utilise une succession de problèmes de discrimination entre deux objets, l'un des deux étant toujours récompensé, le choix effectué à la première présentation d'une nouvelle paire d'objets ne peut être qu'aléatoire. Pourtant, muni des informations collectées grâce à ce premier essai, le sujet possède désormais toutes les informations nécessaires pour résoudre le problème de façon correcte dès le deuxième essai. Dans ce type de tâche, tous les singes se montrent capables d'apprendre la nature même de la tâche. De plus, ces capacités ne semblent pas varier de façon importante des prosimiens aux anthropoïdes. Bien qu'un certain apprentissage puisse être visible dès le 10^e problème, il faut en moyenne 200 problèmes pour que des singes aussi différents que le lémur, l'atèle, diverses espèces de macaques et de grands singes répondent correctement dès le 2^e essai. L'évidence d'une réelle gradation dans la qualité d'un tel apprentissage "conceptuel" le long de l'échelle phylogénétique est cependant obtenue lorsque l'on utilise des tâches de discrimination suivies d'une inversion. Face à une paire d'objets A et B, l'objet A ayant toujours été récompensé, l'inversion consiste à soudainement associer la récompense à l'objet B. Si le singe a appris une association entre stimulus A et récompense, ce premier apprentissage sera néfaste puisqu'il lui faudra passer par une phase d'extinction avant de débiter l'apprentissage de son nouveau problème. Si au contraire l'animal a appris la nature même de la tâche, l'apprentissage du premier problème lui aura été bénéfique et il sera plus rapide à apprendre le suivant. Cette notion "d'index de transfert" a été proposée par Rumbaugh (1970) ; et la synthèse des études qu'il a effectuée sur plusieurs espèces de singes (Rumbaugh et Pate, 1984a, 1984b) montre que cet index de transfert augmente des prosimiens aux singes puis aux anthropoïdes. Alors que les singes du nouveau monde basent plutôt leur décision sur un apprentissage "stimulus-réponse", les singes de l'ancien monde, et tout particulièrement le singe vervet et le macaque rhésus, apprennent la stratégie qui leur permet de résoudre la tâche, faisant preuve de capacités très proches de celles des anthropoïdes.

La plupart des singes de l'ancien monde paraissent donc capables d'acquérir des stratégies et de les utiliser dans des situations nouvelles et

diverses. Encore doivent-ils avoir la possibilité de conserver en mémoire une certaine représentation d'un objet donné, même lorsque cet objet n'est pas visuellement présent dans son environnement. La réalité de cette trace mnésique a notamment été montrée à l'aide de diverses tâches d'appariement différé, dans lesquelles un délai variable est imposé au singe entre la présentation de la situation test et la réponse qu'il doit donner. Les grands singes et les singes de l'ancien monde ont de très bonnes performances dans ce type de tâches (Fischer et Kitchener, 1962 ; Medin, 1969 ; Vauclair, Rollins et Nadler, 1983). Par contre, et même si le capucin a parfois fait preuve d'une bonne rétention (D'Amato et Worsham, 1972) cette capacité apparaît moins développée et moins robuste chez les singes du nouveau monde. Il est important de souligner la souplesse que doit posséder la représentation de l'objet car les grands singes et les macaques rhésus sont capables de réaliser des discriminations cross-modalitaires. Ils peuvent ainsi reconnaître visuellement un objet qu'ils n'avaient eu la possibilité d'explorer précédemment que de façon tactile (Davenport et Rogers, 1970 ; Elliot, 1977 ; Weiskrantz et Cowey, 1975). La première représentation formée est donc basée sur des informations somesthésiques et doit, lors de la deuxième présentation de l'objet, être activée par des informations visuelles. Les représentations mentales utilisées semblent donc suffisamment abstraites pour permettre aux singes de combiner des informations obtenues par différents canaux sensoriels.

Ce résumé très succinct montre cependant que les primates non humains sont capables d'un apprentissage conceptuel et qu'ils peuvent utiliser des représentations mentales présentant un certain degré d'abstraction. Certaines opérations – telles que des rotations mentales – peuvent d'ailleurs être effectuées sur ces représentations. Dans un test d'appariement différé (Hopkins, Fagot et Vauclair, 1993), les babouins sont capables de reconnaître un objet cible par rapport à son image en miroir, même si cet objet a subi une rotation de 60 à 300°. Comme pour l'homme face au même type de tâche, le temps de réaction est corrélé avec l'ampleur de la rotation mentale à effectuer.

Si le singe possède la représentation mentale d'un objet, peut-il posséder la représentation mentale d'une classe d'objets ? Grouper divers éléments – parfois très dissemblables – au sein d'une même classe (généralement définie par l'homme) est le problème de la catégorisation. Qu'en est-il des capacités de catégorisation des singes ? A quel niveau d'abstraction sont-ils capables de catégoriser ? Quels sont les mécanismes cérébraux impliqués ? Existe-t-il des similitudes avec l'homme ?



CATÉGORISATION RAPIDE DE SCÈNES NATURELLES

Les singes paraissent capables de classer des photographies sur la base de la présence ou de l'absence d'un être humain, d'un singe, ou d'un A (D'Amato et Van Sant, 1988 ; Schrier et al., 1984 ; Yoshikubo, 1985). Ils pourraient même manier des concepts beaucoup plus abstraits ; ainsi le singe de Java utiliserait un concept d'affiliation sociale, reconnaissant parmi toutes les photographies de paires de congénères familiers celles qui présentent deux singes liés par la relation mère-enfant (Dasser, 1988). Cette relation est d'ailleurs biologiquement importante puisque, dans certains groupes sociaux, la fille hérite du statut de dominance maternelle. Chez le chimpanzé, la reconnaissance d'une relation mère-fils entre deux individus peut même être effectuée alors que les individus sont totalement inconnus au sujet, montrant, dans cette espèce, une capacité à évaluer les similarités phénotypiques de deux individus (Parr et de Waal, 1999). Mais divers auteurs mettent en doute la capacité des singes à développer des représentations abstraites et favorisent souvent l'hypothèse selon laquelle leurs compétences seraient basées sur la détection de petits ensembles de caractéristiques dont la pertinence aurait été déterminée par l'apprentissage. Si tel était le cas, les singes devraient facilement classer dans une même catégorie des exemplaires très similaires mais devraient présenter de plus en plus de difficultés lorsque la catégorie se complexifie et regroupe des exemplaires de plus en plus différents. Une étude de Roberts et Mazmanian (1988) a comparé l'homme et le singe écureuil dans 3 tâches de catégorisation de différentes complexités. Les sujets étaient placés face à une paire de stimulus, l'un des deux appartenant à une catégorie donnée. Ils étaient récompensés pour choisir ce stimulus. Au niveau le plus concret, les sujets devaient choisir le martin pêcheur, le stimulus non récompensé (distracteur) étant un autre oiseau. Au niveau intermédiaire, ils devaient choisir l'oiseau, le distracteur étant un autre animal, enfin au niveau le plus abstrait, ils devaient choisir l'animal, les distracteurs étant des photographies diverses ne contenant pas d'animaux. Les hommes apprennent et transfèrent sans difficulté leur apprentissage sur un nouvel ensemble de photographies quelle que soit la tâche, la tâche la plus concrète apparaissant même la plus difficile. En revanche, les singes écureuils apprennent avec plus de difficultés et, dans les expériences de transfert sur un ensemble de nouvelles diapositives, ils n'atteignent 80% de réussite que sur la tâche la plus concrète (martin pêcheur versus autres oiseaux). Dans cette tâche, les vives couleurs du martin pêcheur sont d'ailleurs un critère de choix à la fois pour les hommes

et les singes puisque la précision de leur catégorisation chute lorsqu'ils doivent choisir le martin pêcheur parmi d'autres espèces d'oiseaux colorés. L'acquisition de concepts chez l'animal pourrait être ainsi limitée à des catégories concrètes. Pourtant, cette étude présente une autre expérience qui montre que les singes écureuil augmentent très nettement la précision de leur transfert dans la catégorisation la plus abstraite (animal versus non animal) dès qu'on augmente le nombre d'exemplaires qui constituent leur base d'apprentissage. Les concepts les plus abstraits seraient donc simplement plus longs à se constituer, ce qui en soi ne semble pas très étonnant.

Au laboratoire, nous avons repris l'étude de la catégorisation la plus abstraite (animal versus non animal) chez le macaque rhésus et nous avons obtenu un bien meilleur taux de transfert (Fabre-Thorpe, Richard et Thorpe, 1998) que celui atteint par le singe écureuil. Plusieurs points différencient notre tâche de celle de Roberts et Mazmanian. Dans cette dernière, les stimulus sont toujours présentés par paires, l'un des deux étant récompensé ; l'animal peut donc regrouper les exemplaires sous une étiquette étroitement associée à la récompense qu'il reçoit. Notre tâche fait appel à une catégorisation go/no-go dans laquelle les images sont présentées séquentiellement. Les cibles étant associées à des réponses motrices ("go") et les distracteurs étant associés à l'absence de cette même réponse ("no go"), chaque réponse correcte est récompensée. L'animal ne peut donc pas constituer deux ensembles de stimulus basés sur leur association (ou non) avec une récompense. De plus, la base d'apprentissage des sujets est chaque jour agrandie et renouvelée : l'animal est confronté quotidiennement et depuis le début de son apprentissage à de nouvelles images qui apparaissent au hasard parmi des stimulus appartenant à sa base d'apprentissage. Cette procédure rend peu efficace la mise en place d'une stratégie basée sur une association stimulus-réponse. Elle assure une stabilité des niveaux d'attention et de motivation grâce aux performances de catégorisation des images familières et permet un suivi longitudinal des capacités de transfert sur les images nouvelles. Enfin, nous avons utilisé des temps de présentation extrêmement courts, les stimulus ne sont flashés que pendant 30 à 80 ms et, lorsqu'une cible est détectée, le singe doit effectuer sa réponse motrice en moins de 1 s. Notre tâche place donc le système visuo-moteur dans des conditions extrêmes de fonctionnement, puisque la brièveté de la présentation de l'image ne permet pas au sujet d'effectuer de mouvement oculaire. Cette contrainte oblige le sujet à baser sa décision sur des mécanismes de catégorisation rapides, qui se sont d'ailleurs révélés d'une extrême efficacité. Pour apprécier les similitudes et les dissemblances de ces processus chez l'homme et le singe, cette étude a été conduite parallèlement chez le rhésus

et chez des sujets humains qui ont été testés dans le même dispositif, avec la même tâche et les mêmes stimulus.

La tâche

Deux tâches de catégorisation visuelle avec réponse go/no-go ont été étudiées, l'une basée sur la classification des aliments (vs. non aliment – Al/NAI) et l'autre sur la classification des animaux (vs. non animal – An/NAn). Les sujets (deux singes rhésus et deux groupes de 10 hommes) sont placés face à un écran tactile et déclenchent eux-mêmes la présentation séquentielle des photographies en posant une main sur une touche tactile placée sous l'écran (Figure 1). Chaque image représente une scène naturelle contenant (ou non) une cible ; elle est flashée centralement pendant un temps extrêmement bref de moins d'un dixième de seconde (de 30 à 80 ms). Si l'image contient une cible (aliment ou animal, suivant la tâche) le sujet doit, en moins d'une seconde, relâcher la touche tactile et appuyer sur l'écran à l'endroit où la diapositive s'était affichée. Dans le cas contraire, il doit maintenir l'appui sur la touche tactile pendant au moins 1 s. Chaque réponse correcte (go après une cible ou no-go après un distracteur) est signalée par un son et, pour le singe, récompensée par une gorgée de jus de fruit ; en revanche chaque erreur est punie par le réaffichage de l'image incorrectement classée pendant 3 s sur l'écran.

Les sujets doivent prendre une succession de décisions rapides basées sur des images présentées pendant un temps si bref que tout mouvement oculaire d'exploration est impossible. De plus, ils ne possèdent aucune information sur le stimulus qui va leur être proposé. Cibles et distracteurs apparaissent au hasard, de façon équiprobable. Toutes les photographies sont des scènes naturelles provenant d'une grande banque de données commerciales (Corel) disponible sur CD ROM. Les cibles sont extrêmement variées et présentées dans un environnement naturel (Figure 2) : mammifères, oiseaux, reptiles, poissons pour les animaux ; fruits, légumes, salades, gâteaux pour les aliments. Les distracteurs incluent des paysages, des arbres, des fleurs, des monuments, des objets manufacturés ainsi que les objets cibles de l'autre catégorie. De plus, cibles et distracteurs font alterner gros plans et vues générales. Lorsqu'il s'agit d'une cible, le sujet n'a pas d'information sur son type (parmi le large éventail possible), sa taille, ou sa position dans l'image. La cible peut être partiellement cachée par un autre objet présent dans la scène ou présentée seulement en partie, mais plusieurs cibles peuvent également être présentes. Avec une telle variété de

stimulus, l'aide du contexte peut ainsi être considérée comme nulle ou négligeable.

Figure 1. Le sujet (A) – homme ou singe –, est assis face à un écran tactile (B). Pour démarrer la séquence d'images, il appuie sur une touche tactile (C) située sous l'écran. Le sujet doit maintenir cet appui si un distracteur est présenté, en revanche, dès qu'une cible est détectée (un animal par exemple), il doit relâcher son appui et toucher l'écran à l'endroit où la photographie a été présentée. La séquence recommence lorsqu'il appuie de nouveau sur la touche.

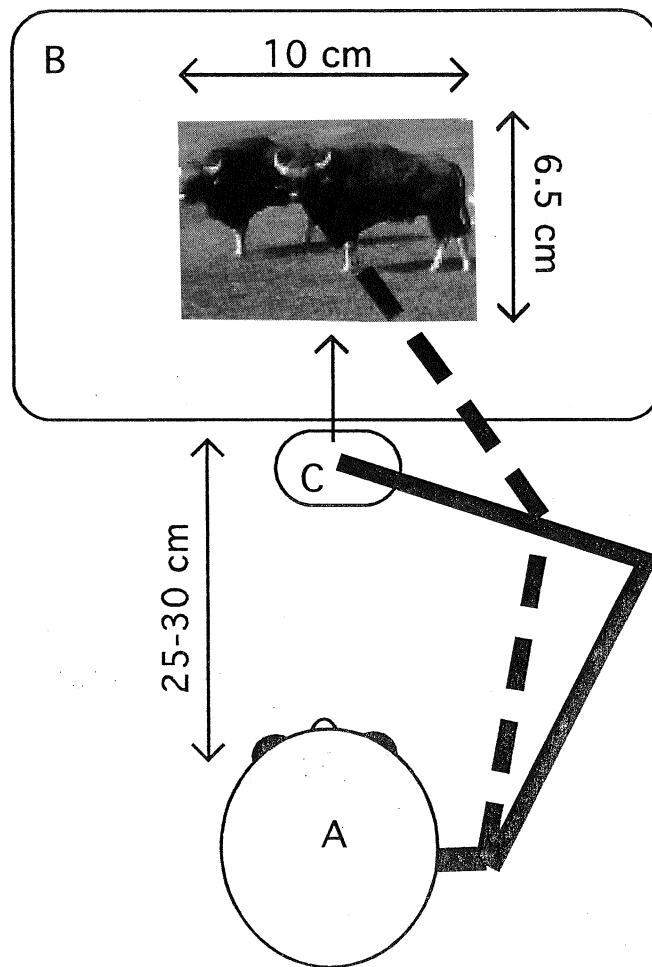


Figure 1. Task and set-up.

Figure 2. Pour illustrer la diversité des images naturelles proposées et les similarités possibles entre cibles et distracteurs, cette figure présente (pour chacune des deux catégorisations étudiées), six cibles (colonnes de gauche) et six distracteurs (colonnes de droite).

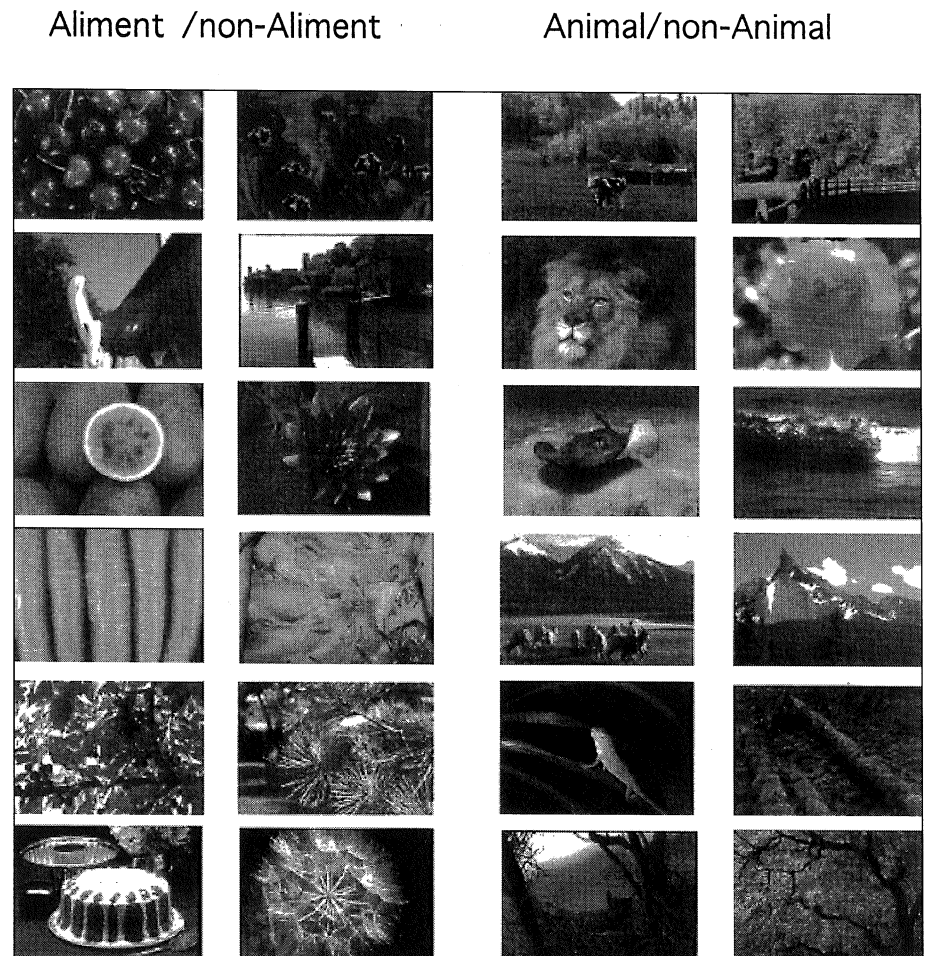


Figure 2. Example of natural stimuli used as targets or distractors.

Chaque sujet humain travaille sur des séries de 100 images qu'il n'a jamais vues auparavant. Chaque singe travaille quotidiennement sur une série de 100 images qui peuvent lui être présentées plusieurs fois, mais toujours dans un ordre différent pour éviter tout apprentissage de séquence.

Parmi ces 100 stimulus, 10 n'ont jamais été vus par l'animal (5 cibles et 5 distracteurs). Ce protocole permet donc d'analyser tout à la fois (1) la réponse du singe à la première présentation d'un stimulus nouveau (10 essais par jour), tâche équivalente à la catégorisation effectuée par les sujets humains, (2) son apprentissage sur ce stimulus, et (3) ses performances lorsque ce stimulus est devenu très familier. Hommes et singes travaillent sur le même dispositif, et sont placés face aux mêmes diapositives. Leurs capacités de catégorisation peuvent ainsi être comparées. De plus l'effet perturbateur de diverses manipulations des stimulus sur leur niveau de performance peut également être analysé et donner lieu à une étude comparative. L'analyse des performances porte sur le taux de réponses correctes et sur la vitesse d'exécution. Cette vitesse est évaluée par le temps de réaction (délai séparant le début de la présentation de l'image sur l'écran et l'instant où la main quitte l'appui sur la touche). Les erreurs de catégorisation des hommes et des singes font également l'objet d'une étude comparative précise.

Catégorisation chez le macaque rhésus

Les deux macaques ont été entraînés à la tâche de catégorisation go/no-go sur un petit ensemble de 10 photographies auxquelles de nouveaux stimulus (2, 4, 6, puis 10) étaient ajoutés quotidiennement. Dès que l'ensemble d'images connues était suffisant, les 10 nouvelles photographies étaient présentées au hasard parmi 90 images "familières". Les animaux se sont montrés extrêmement rapides à généraliser leur apprentissage aux photographies qu'ils n'avaient jamais vues et, dès que la précision de leur catégorisation et leur vitesse de réponse se sont stabilisées, leurs performances ont été évaluées sur un premier groupe de 200 nouvelles photographies (100 cibles et 100 distracteurs). L'ensemble de stimulus déjà vus atteignait alors 280 images pour la catégorisation AI/NAI et 340 pour la catégorisation An/NAn.

Sur ce premier groupe de 200 nouveaux stimulus, il faut tout d'abord noter que, malgré le défi que constitue cette tâche, les performances obtenues ont été excellentes. Le pourcentage de catégorisations correctes a atteint 90.5% sur la tâche AI/NAI et 84% dans l'autre cas. De plus, les singes se sont révélés particulièrement rapides à déclencher leur réponse puisque leurs temps de réaction médians étaient respectivement de 344 et de 245 ms.

Figure 3. Distribution des temps de réaction (en millisecondes) des réponses déclenchées par les singes (en haut) et par les dix sujets humains (en bas) lors de la classification Animal/non-Animal d'un total de 400 images en couleur (deux ensembles de 200 images – voir texte). Les réponses correctes déclenchées en direction des cibles (en évidé) et les réponses incorrectes déclenchées en direction des distracteurs (en noir) sont présentées en pourcentage du nombre total des réponses produites par pas de temps de 10 ms. Il faut noter que le singe est beaucoup plus rapide que l'homme.

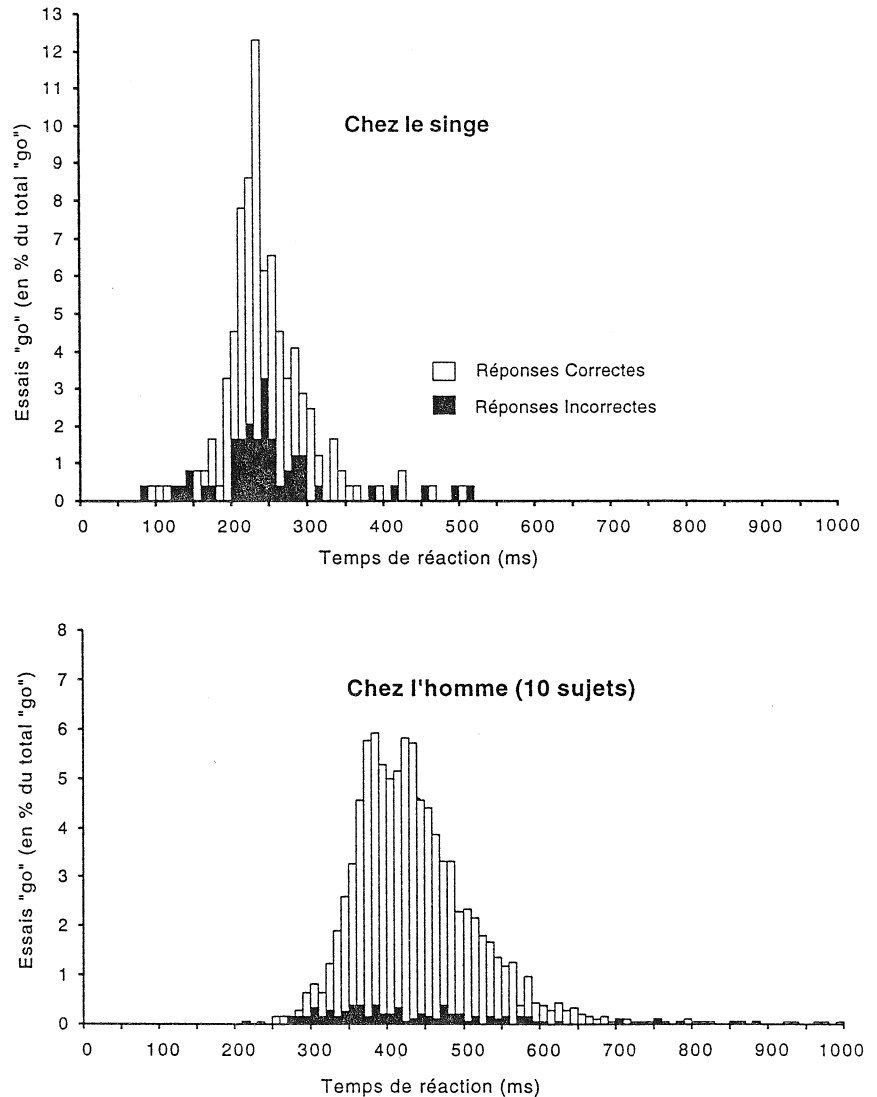


Figure 3. Reaction time distributions for correct and incorrect go responses in monkeys and humans.

L'expérience a été renouvelée avec, pour chaque animal, un nouvel ensemble de 200 nouvelles photographies. Les performances des deux macaques ont été quasi identiques en ce qui concerne le pourcentage de catégorisations correctes, 89.5% sur la tâche AI/NAI et 87% dans l'autre cas. Les singes se sont même révélés plus rapides que pour le premier groupe d'images (temps de réaction médian de 311 ms sur la tâche AI/NAI et de 233 ms dans l'autre cas), effet probablement dû à l'apprentissage. La distribution des temps de réaction est présentée pour l'ensemble de ces deux groupes d'images pour le singe travaillant sur la catégorisation An/Nan (Figure 3) et permet de montrer que les réponses "go" correctes sont statistiquement plus nombreuses que les réponses "go" incorrectes dès 180-190 ms.

Comparaison avec l'homme

Pour l'étude comparative, ces deux ensembles de 200 nouvelles images ont donc été proposés aux hommes. Le taux de réponses correctes atteint 95% en moyenne chez le sujet humain et les réponses sont produites avec un TR médian variant de 383 ms à 504 ms. Pour chacun des ensembles d'images, les hommes se sont toujours révélés plus précis mais beaucoup plus lents que les singes (Figure 3).

Hommes et singes sont donc capables d'excellentes performances dans ce type de tâche. Pour aller plus avant dans cette étude comparative, les images qui avaient induit des erreurs de catégorisation ont été tout particulièrement analysées.

Les réponses erronées des singes étaient souvent déclenchées en réponse aux images qui produisaient également des erreurs de catégorisation chez les hommes ; pour tenter de quantifier objectivement ce résultat, nous avons classé les images qui avaient été proposées en fonction du degré de difficulté qu'elles avaient présenté pour l'homme. Chaque groupe de 200 photographies a donc été scindé en 3 sous-ensembles : les images faciles (les 10 sujets humains avaient tous fourni la bonne réponse), les images de difficulté intermédiaire (une réponse erronée avait été fournie par 1, 2 ou 3 sujets humains) et les images très difficiles (4 sujets humains au moins s'étaient trompés dans leur catégorisation). Les performances du singe ont ensuite été ré-analysées séparément dans chacun de ces trois sous-ensembles. Les deux animaux avaient catégorisé avec un pourcentage élevé de réussite (93% en moyenne) les images faciles, ce taux de réussite était plus bas pour les images moyennement difficiles (78% en moyenne) et

Figure 4. Précision du singe pour chacune des catégorisations (Animal/non-Animal et Aliment/non-Aliment) et pour chacun des trois groupes d'images (Faciles, Intermédiaires et Difficiles) constitués à partir des performances des 10 sujets humains. Les deux colonnes en trame pointillée représentent la précision atteinte pour les deux ensembles de 200 nouvelles images présentées en couleur. Les deux autres colonnes concernent l'ensemble de 400 nouvelles images dont 200 étaient présentées en couleur (trame pointillée serrée) mais mélangées au hasard parmi 200 images présentées en noir et blanc (trame noire). Il faut noter que, pour les stimulus chromatiques, le singe est très précis sur les images faciles, moins précis sur les images intermédiaires et au niveau de la chance pour les images difficiles. Noter également que cette correspondance est moins bonne pour les images en noir et blanc (voir texte).

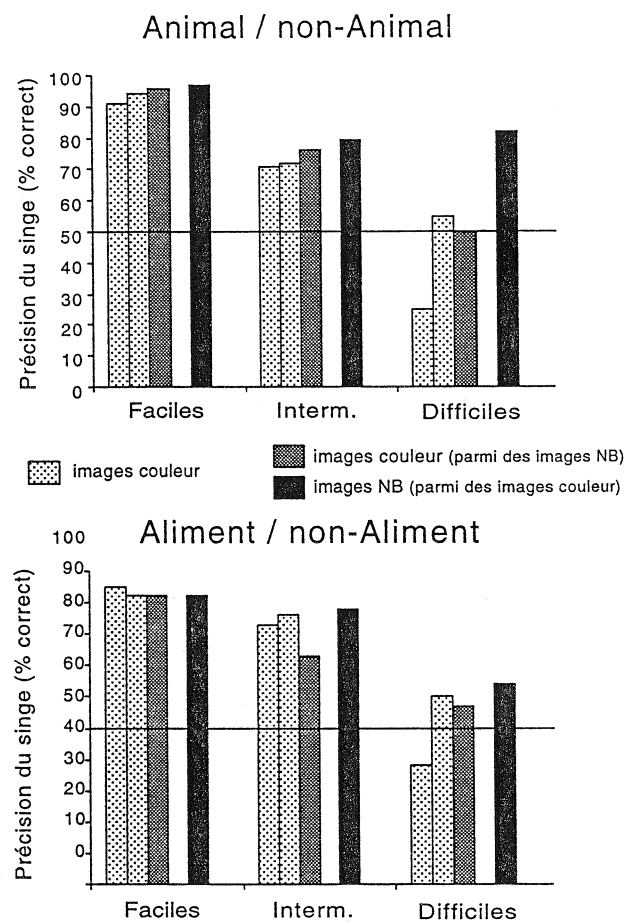


Figure 4. Correlation between monkey and human categorisation accuracy.

chutait au niveau de la chance pour les images très difficiles (44.5% en moyenne). Cet effet est très robuste puisqu'il est apparu pour chaque singe et pour chaque ensemble d'images étudié (Figure 4).

Ces données soulèvent deux points importants. L'un concerne les similarités étonnantes constatées dans les performances de catégorisation de l'homme et du singe et pose le problème de leurs limites. L'autre concerne la rapidité des réponses enregistrées, rapidité qui impose une contrainte temporelle forte sur le type de traitement de l'information visuelle possible dans un délai aussi court.

SIMILARITÉS DES PERFORMANCES ENTRE L'HOMME ET LE SINGE

Les analogies enregistrées dans les performances et les erreurs de catégorisation de l'homme et du singe permettent de soulever la question de la similitude des traitements visuels à la base du comportement de classification rapide dans les deux espèces. Si les mécanismes cérébraux sont comparables et basés sur les mêmes indices, des manipulations expérimentales effectuées sur les images proposées devraient affecter les performances de l'homme et du singe de façon similaire. Nous avons étudié deux cas précis :

Comment sont affectées les performances lorsque les informations chromatiques sont absentes, l'image étant présentée en noir et blanc ?

Comment sont affectées les performances lorsque les images sont devenues familières et que des caractéristiques visuelles de "bas niveau" pourraient être suffisantes pour assurer une bonne classification ?

Catégorisation en noir et blanc

Parmi les indices simples qu'hommes et singes pourraient utiliser pour identifier les aliments et les animaux, la couleur pourrait jouer un rôle important. La couleur est un important indice pour la catégorisation du martin pêcheur dans l'étude de Roberts et Mazmanian (1988). Elle semble jouer un rôle dans la capacité du capucin à catégoriser des photographies sur la base de la présence ou de l'absence d'êtres humains ; d'Amato et Van Sant (1988) soulignent que ces singes pourraient avoir pris leur décision sur la base d'indices simples comme la présence d'une tâche de couleur rouge. Pourtant l'importance de la couleur dans les performances de catégorisation des animaux reste controversée ; pour certains auteurs, la précision de caté-

gorisation des pigeons reste inchangée lorsque les images sont présentées en noir et blanc (Herrnstein et Loveland, 1964), alors qu'elle est perturbée pour d'autres auteurs (Aust et al., 1999) et peut chuter jusqu'au niveau de la chance pour la catégorisation "aliments" versus "non aliments" (Watanabe, 1997). Le rôle joué par les indices de couleur reste donc incertain, et pourrait varier avec le type de catégorisation étudiée ; de plus il faut tenir compte de l'espèce animale utilisée, puisque les résultats obtenus chez l'oiseau, dont la vision est pentachromatique, ne peuvent pas être étendus aux singes de l'ancien monde, dont la vision trichromatique est proche de celle de l'homme (Dulai, Bowmaker, Mollon et Hunt, 1994).

A l'aide d'un ensemble de 200 images familières, les deux macaques ont été entraînés à catégoriser indifféremment des photographies présentées au hasard soit en couleur soit en noir et blanc. Leurs performances ont ensuite été étudiées sur un nouvel ensemble de 400 images (200 en couleur et 200 en noir et blanc) qu'ils n'avaient jamais vues auparavant. Comme dans la première expérience, les images nouvelles apparaissaient parmi des images familières. Chaque jour, 20 nouvelles images étaient introduites (10 en couleur et 10 en noir et blanc ; cibles et distracteurs étant toujours équiprobables dans chacune des modalités).

Si une analyse fine a permis de faire apparaître certaines perturbations dans la catégorisation des aliments (Delorme, Richard et Fabre-Thorpe, soumis), les performances de catégorisation des singes (en précision et en rapidité) se sont révélées globalement similaires (Figure 5), quelles que soient les conditions de présentation, couleur ou noir et blanc (NB). Le pourcentage de réponses correctes pour la classification des aliments a atteint 85.5% en couleur et 87.5% en NB, alors que ces valeurs étaient respectivement de 90.5 et de 91.5% pour la classification des animaux. Dans les deux cas, les distributions des temps de réaction des réponses "go" correctes sont restées inchangées. Les TR médians dans les conditions couleur et NB sont respectivement de 297 ms et 306 ms pour les aliments et de 266 ms et 269 ms pour les animaux (Test bilatéral U de Mann Whitney, ns).

Deux groupes de 10 sujets humains ont été testés sur ces mêmes 400 photographies (Figure 5). En moyenne, la précision globale des sujets humains chute de 2% lorsque les stimulus sont présentés en NB (93.2% en couleur et 91% en NB). Cette tendance devient significative pour la catégorisation "animal" (95.7% de réponses correctes pour les images en couleur et 93% pour les images NB, χ^2 bilatéral = 13.6, d.l. = 1, $p = 0.0002$) et pour la détection des cibles dans la catégorisation "aliments". En ce qui concerne la vitesse de décision elle n'est que peu (ou pas) affectée. Le

temps de réaction est inchangé pour la catégorisation An /nAn (TR médian : 412 ms en couleur, 415 ms en NB), et il n'est allongé que d'une dizaine de millisecondes pour la détection des aliments (TR médian : 427 ms en couleur, 439 ms en NB ; test U de Mann Whitney, $p < 0.0001$).

Figure 5. Comparaison, pour chacune des catégorisations étudiées, des performances moyennes enregistrées chez le singe et chez l'homme en fonction de la présence ou de l'absence des indices de couleur dans les scènes naturelles proposées. Pour les images en couleur (trame noire) et pour les images en noir et blanc (trame hachurée), la précision de la catégorisation (en haut) est exprimée en pourcentage de réponses correctes ; et la rapidité d'exécution (en bas) est illustrée par le temps de réaction médian (en millisecondes) des réponses motrices correctes déclenchées en direction des cibles.

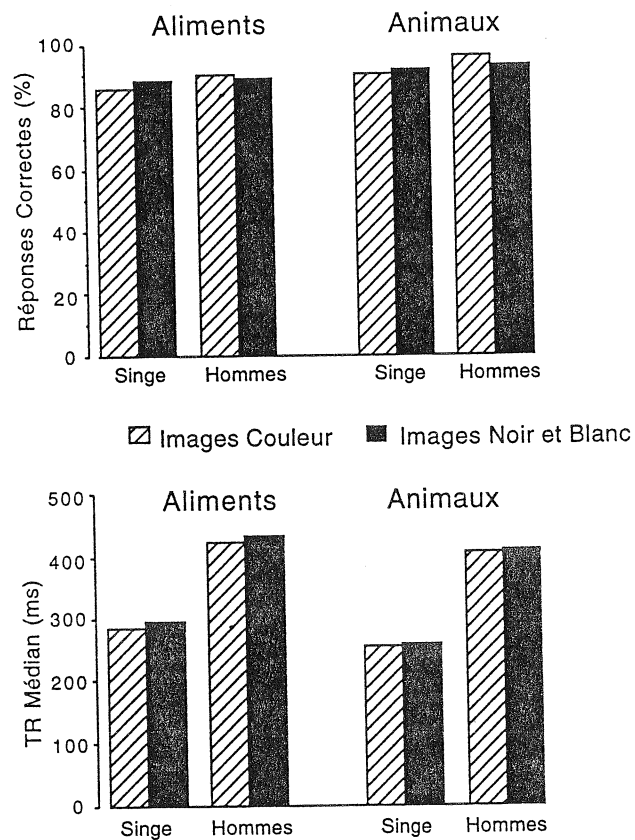


Figure 5. Accuracy and speed of response in the categorisation task with chromatic and achromatic stimuli.

Cependant l'importance des indices de couleur était très différente d'un individu à l'autre, la précision de certains hommes restant totalement inchangée malgré l'absence de couleur alors que pour d'autres elle baissait de 6%. En fait, une corrélation importante apparaît entre la vitesse de réaction d'un individu et la détérioration de sa performance en l'absence de couleur. La précision des sujets les plus rapides reste inchangée mais plus le sujet est lent à produire sa réponse, plus sa précision est affectée par l'absence de couleur (Figure 6).

La couleur ne serait donc un indice pertinent que pour les individus les plus lents à déclencher leur réponse. Dans ce type de catégorisation, les singes se comporteraient comme les sujets humains les plus rapides, associant vitesse de décision et indifférence aux indices de couleur. La comparaison des erreurs commises par les singes et les hommes sur le groupe des 200 images en couleur montre, comme précédemment, que les difficultés rencontrées par les deux espèces sont similaires. Lorsque les images sont présentées en NB, cette corrélation, encore présente, est moins bonne, reflétant ainsi la prise en considération des indices de couleur par certains des sujets humains (Figure 4). Les sujets les plus rapides pourraient baser leur catégorisation sur les informations visuelles les plus précocement disponibles, à savoir les informations achromatiques transmises par la voie magnocellulaire ; en revanche la décision des sujets les plus lents prendrait également en compte les caractéristiques chromatiques des stimulus, et ferait alors intervenir la voie visuelle parvocellulaire, plus lente dans le transfert d'informations.

Catégorisation d'images familières

Les singes travaillant chaque jour sur de nombreux essais, il a été possible de comparer la façon dont leurs performances se modifiaient lorsque le même stimulus leur était présenté plusieurs fois. Chez l'homme, une série expérimentale a été menée dans un dispositif quasi-identique (le sujet devait seulement relever le doigt de la touche tactile pour signaler la présence d'une cible). Chaque sujet humain se familiarisait chaque jour, pendant trois semaines, avec un ensemble de 200 images. Ces images "familières" étaient ensuite présentées mélangées à des images totalement nouvelles. Pour le singe comme pour l'homme, une plus grande familiarité avec un stimulus permet une augmentation de la précision surtout dans la détection des cibles et une diminution des temps de réaction pour les images contenant des cibles particulièrement difficiles à repérer. Cependant, en ce qui

Figure 6. Corrélation obtenue chez l'homme entre temps de réaction et précision des catégorisations en fonction de la disponibilité des indices de couleur. Pour chacune des deux catégorisations étudiées, chacun des 10 sujets humains est représenté par une astérisque avec en abscisse son temps de réaction médian (exprimé en millisecondes), et en ordonnée la différence de précision (exprimé en pourcentage de réponses correctes), calculée en retranchant le pourcentage de réponses correctes enregistré pour des stimulus en noir et blanc du pourcentage correct enregistré pour des stimulus chromatiques. Cette corrélation montre que les indices de couleur ne sont quasiment pas utilisés par les sujets les plus rapides.

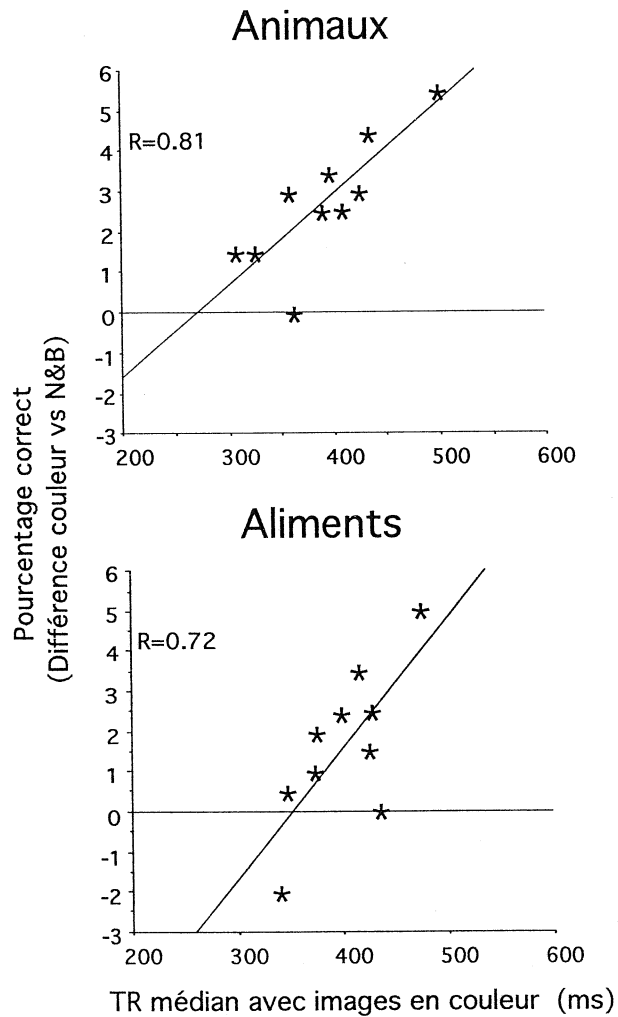


Figure 6. Speed of response and use of chromatic cues.

concerne la grande majorité des images, les temps de réaction restent incompressibles montrant que, dans ce type de catégorisation visuelle ultra-rapide, la stratégie de classification ne se modifie pas avec l'apprentissage.

PROCESSUS NEURONAUX IMPLIQUÉS

L'une des principales caractéristiques de l'expérience que nous avons conduite est l'utilisation de stimulus flashés pendant un temps très court. Pratiquement toutes les études de catégorisation effectuées chez l'animal utilisent des stimulus qui sont affichés pendant des temps longs (voire très longs), variant de 4 à 30 s. Il se pourrait que de telles durées de présentation permettent la mise en place de stratégies variées incluant des explorations oculaires précises et l'utilisation de processus complexes de vérification qui accentuent les différences entre hommes et singes. En revanche, les présentations brèves que nous utilisons encouragent nos sujets (humains ou non humains) à des décisions rapides et instinctives basées sur des traitements rapides de l'information visuelle. Dans ces conditions, les performances obtenues sont encore excellentes ; elles sont surtout remarquablement similaires chez l'homme et le singe.

Singes et hommes sont donc capables de catégoriser avec une grande précision et une grande rapidité des scènes naturelles qu'ils n'ont jamais vues auparavant et qui ne leur sont présentées que pendant 30 à 80 ms ; l'homme restant plus précis que le singe et le singe plus rapide que l'homme. L'étude des erreurs de classification montre qu'il existe une grande superposition des ensembles d'images qui induisent des erreurs chez les hommes et les singes. Cette superposition reflète ainsi la similitude des difficultés qu'ils rencontrent. De plus, leurs performances sont affectées de la même façon, (i) par l'entraînement (lorsqu'un même stimulus a été présenté et catégorisé de nombreuses fois), et (ii) par la suppression des indices de couleur. Comment ces données peuvent-elles nous renseigner sur les processus cérébraux impliqués ?

A partir des aires visuelles occipitales, deux principales voies de traitement des informations visuelles sont distinguées : une voie dorsale qui jouerait un rôle dans la perception de l'espace, les relations spatiales entre objets et l'action dirigée vers l'objet, et une voie ventrale dont le rôle serait prépondérant dans l'identification des formes (Milner et Goodale, 1993 ; Ungerleider et Mishkin, 1982). L'anatomie et la physiologie de ces deux voies ont fait l'objet de très nombreuses études chez le singe et la connais-

sance fine ainsi disponible sur les circuits visuels du primate non-humain est, par rapport à l'homme, un atout considérable.

Cette voie ventrale se dirige vers le cortex inféro-temporal (IT) dans lequel la réponse des cellules peut être liée à la présentation visuelle d'objets spécifiques comme des visages (Perrett, Rolls et Caan, 1982) ou d'autres objets familiers ou expérimentaux (Booth et Rolls, 1998 ; Logothetis, Pauls et Poggio, 1995). Cette réponse est robuste à une grande diversité de présentation de l'objet, images colorées, dessins, silhouettes, contours illusoire (Chadaide, Kovács, Köteles et Benedek, 1999). Si l'on considère un objet donné, l'une des catégorisations les plus élémentaires consiste à classer, sous la même étiquette, toutes les vues sous lesquelles il peut être présenté. Les cellules de IT sont très souvent sélectivement liées à la présentation de deux, voire trois vues du même objet. Mais leur réponse peut aussi, dans certains cas, être totalement "vue-invariante" ; dans ce cas, la cellule répond quel que soit l'angle de vue dans lequel l'objet lui a été présenté et cette réponse résiste à des modifications de taille ou de localisation de l'objet. Cette propriété a été montrée chez le macaque pour des objets nouveaux qui avaient été introduits dans leur cage quelques semaines avant que ne soient effectués les enregistrements (Booth et Rolls, 1998). Les cellules de IT peuvent également répondre sélectivement à plusieurs objets qui appartiennent à la même classe ; chez le singe catégorisant des photographies d'arbres, un quart des cellules répondent différemment aux arbres et aux non-arbres (Vogels, 1999). Et bien que Booth et Rolls ne l'aient pas interprété ainsi, cette sélectivité à une classe d'objets pourrait avoir été observée dans leur étude puisque certaines des cellules enregistrées étaient sélectives à plusieurs des objets qu'ils avaient laissés dans la cage de leurs macaques. Le macaque aurait pu spontanément créer une classe regroupant les objets dont il avait la libre utilisation. Mais la réponse de ces neurones ne code pas tous les exemplaires d'une classe, elle ne code qu'un certain nombre d'exemplaires. La catégorie "arbre" n'est pas représentée dans sa totalité au niveau d'un seul neurone. Ces données montrent que, si certaines cellules peuvent coder plusieurs exemplaires d'une même catégorie et, ceci, quel que soit leur angle de présentation, elles ne sont pas suffisamment "stimulus-invariant" pour coder tous les exemplaires entrant dans cette catégorie. Ces données électrophysiologiques très récentes confortent plutôt l'idée de catégories basées sur de multiples prototypes – chacun autorisant un certain degré de généralisation – et ceci par opposition à la théorie du prototype unique. Une catégorie serait alors représentée par l'activité d'une population de neurones ; chacun de ces neurones codant des ensembles d'exemplaires différents, se chevauchant à des degrés divers.

Il existe donc de forts arguments pour penser que c'est au niveau du cortex IT que s'effectuerait l'identification d'un objet et la reconnaissance de son appartenance à une catégorie donnée. Or, les réponses de ces cellules peuvent apparaître à des latences très brèves 80-100 ms (Oram et Perrett, 1992 ; Perrett et al., 1982 ; Vogels, 1999). La brièveté des temps de réaction comportementaux des singes comme des hommes apparaît donc en accord avec la rapidité de réponse enregistrée au niveau neuronal. Ainsi, si l'on s'intéresse aux réponses "go" déclenchées par le singe avec les latences les plus courtes, le délai minimal permettant un traitement complet des informations visuelles peut être apprécié. Si les réponses correctes et incorrectes précoces ne sont que des anticipations issues d'un traitement inachevé, elles devraient se répartir de façon équivalente puisque distracteurs et cibles sont équiprobables. C'est ce qui est observé (Figure 3) jusqu'à une latence de 170-180 ms, latence à laquelle les réponses deviennent statistiquement biaisées en direction des réponses correctes. Ce résultat est surprenant puisqu'il montre qu'un délai aussi court suffit pour que soient effectuées toutes les étapes entre l'entrée visuelle et la sortie motrice (traitement visuel, prise de décision, déclenchement de l'acte moteur). Ce temps de réaction comportementale incluant une composante motrice non négligeable, le seul traitement visuel ne peut donc pas excéder une centaine de millisecondes. Ces données sont en accord avec une étude effectuée récemment et qui montre, grâce à l'analyse des potentiels évoqués associés à la tâche, que l'information visuelle dans une tâche de catégorisation peut être traitée en 150 ms chez l'homme (Thorpe, Fize et Marlot, 1996). Pourtant, pour transiter de la rétine jusqu'à IT, l'information visuelle doit passer par de nombreuses étapes ; comment les processus cérébraux menant à une catégorisation complexe peuvent-ils se dérouler dans un temps aussi court ?

Les données disponibles dans la littérature permettent d'évaluer, chez le singe, le nombre d'étapes entre la rétine et le cortex temporal où les neurones présentent ces réponses spécifiques. Si un délai d'une centaine de millisecondes permet le traitement de l'information visuelle, cette contrainte de temps montre que le traitement de l'information doit donc être essentiellement unidirectionnel (feed-forward) et massivement parallèle. L'importance et l'efficacité des traitements feed-forward ont déjà été soulevées théoriquement (Thorpe et Imbert, 1989) et renforcées par des études expérimentales menées chez l'homme (Thorpe et al., 1996). Mais il apparaît difficile de réconcilier ces données avec certaines théories du traitement de l'information visuelle impliquant des mécanismes trop lents, comme le développement d'oscillations ou l'utilisation de nombreuses boucles de feed-back. La caractéristique feed-forward de ce traitement est renforcée par le fait que,

dans l'étude présentée ici, ni l'homme ni le singe ne réussissent à produire des réponses à plus courtes latences lorsque les stimulus deviennent hautement familiers. L'enchaînement des processus de traitement semble ainsi ne pas pouvoir se dérouler plus rapidement, comme si, pour les réponses les plus précoces, le nombre d'étapes de traitement était incompressible.

La dernière donnée expérimentale présentée ici concerne les excellentes performances de l'homme et du singe lorsque les scènes naturelles leur sont présentées en noir et blanc, et renforce l'idée d'une identification basée sur des processus feed-forward. En effet, toutes les informations visuelles véhiculées depuis la rétine jusqu'au cortex ne le sont pas à la même vitesse. De récentes données montrent que la voie magnocellulaire est plus rapide que la voie parvocellulaire (Nowak, Munk, Girard et Bullier, 1995). Or les informations de couleur sont véhiculées par la voie parvocellulaire. Tout traitement unidirectionnel basé sur les premières informations visuelles disponibles devrait donc être "aveugle aux couleurs". Nos données renforcent cette hypothèse, puisque les sujets humains dont les performances étaient perturbées par l'absence de couleur étaient également ceux dont les temps de réaction étaient les plus longs. Les singes, comme les sujets humains les plus rapides, présentaient des performances identiques en précision et en vitesse, que les scènes soient proposées en couleur ou en noir et blanc. Ici encore les enregistrements cellulaires effectués dans IT sur le singe éveillé renforcent cette hypothèse. Dans leur grande majorité, les informations de forme sont nécessaires et suffisantes pour obtenir une réponse maximale (Booth et Rolls, 1998 ; Tanaka, Saito, Fukada et Moriya, 1991), seule une minorité de cellules nécessite la combinaison des informations de forme et de couleur.

Si hommes et singes basent cette première identification sur le même type de traitement de l'information visuelle, il reste enfin à expliquer la plus grande rapidité du singe par rapport à l'homme, rapidité qui se retrouve d'ailleurs dans de nombreuses tâches visuo-motrices. En fait, la précision du singe étant inférieure à celle de l'homme, ces observations pourraient n'être que le reflet d'un simple échange précision-vitesse ; mais une seconde explication possible se base sur la différence de taille du cerveau de l'homme et de celui du singe. De récentes données montrent que la vitesse de conduction dans le cortex peut être extrêmement lente, inférieure à 2 m/s et parfois à 0,2 m/s (Frégnac et Bringuier, 1996 ; Nowak et Bullier, 1997) ; des délais considérables pourraient ainsi être dévolus au seul transfert de l'information, délais qui, chez le singe, seraient notablement plus courts puisque les distances à parcourir sont plus courtes.

Singes et hommes sont donc dotés de processus de catégorisation ultrarapides et remarquablement efficaces qui présentent entre eux de grandes similarités. En fait, il ne serait pas étonnant que des milliers d'années d'évolution aient mis en place une capacité de catégorisation rapide et efficace qui présente un réel avantage biologique. Cette capacité serait basée sur un traitement unidirectionnel de la première vague d'informations ; processus qui conduirait à une première identification, correcte dans la grande majorité des cas. Ce premier traitement rapide pourrait orienter les traitements visuels plus tardifs et permettre une perception fine beaucoup plus "intelligente". Cette idée d'une analyse séquentielle, d'abord globale, puis plus fine, est renforcée par d'autres études dont une toute récente, réalisée chez le singe (Sugase, Yamane, Ueno et Kawano, 1999), qui montre que dans le cortex temporal certaines cellules répondant à la présentation de visages produisent une réponse en deux étapes, la première à courte latence indiquerait seulement la présence d'un visage ; la seconde plus tardive (environ 50 ms) contiendrait des informations plus spécifiques, notamment sur l'appartenance du visage à une espèce (primate humain ou non-humain) ou sur l'expression faciale proposée. Ainsi deux types de traitement se succéderaient. Le premier, rapide, massivement parallèle et essentiellement feed-forward, pourrait expliquer la rapidité des identifications effectuées par l'homme dans certains actes de la vie quotidienne lorsque, sans attention particulière, il feuillette rapidement un magazine ou qu'il zappe d'une chaîne de télévision à l'autre ; il pourrait être largement automatique et inconscient. Le second, plus lent, ferait appel à la comparaison de diverses représentations mentales et impliquerait de nombreuses interconnexions entre aires cérébrales liées à la vision, à la mémoire, aux émotions ; il permettrait une identification plus spécifique de cet objet et de toutes les caractéristiques plus ou moins abstraites qui lui ont été associées par apprentissage.

REMERCIEMENTS

Ce travail a reçu l'aide financière du CNRS, de l'Université Paul Sabatier, du GIS "Sciences de la cognition" et du Conseil Régional Midi-Pyrénées.

ABSTRACT

Since the pioneer study by Herrnstein and Loveland (1964) many studies have shown that animals, mainly pigeons and monkeys, could learn to classify a large range of stimuli. But what are the neural mechanisms underlying such ability? Pigeons seem to rely on some simple features including brightness, colour, shading. But, do pigeons, monkeys, or humans classify on the same basis? In the present study we have studied the behaviour of rhesus macaques and human subjects in a rapid categorisation task. Subjects had to respond as quickly as possible when they could detect a given target in a natural scene that they had never seen before. Such rapid categorisation task encourages the subjects to rely on their first decision thus on the first wave of visual processing without involving time consuming processing loops or cognitive operations that would enhance differences in between the strategies used by the two species. Two types of categorisation were studied: "Food/non-Food" and "Animal/non-Animal". All stimuli were natural images from a large commercial data base (Corel). The subject started the sequence of stimuli by pressing on a button, each stimulus was displayed for only 30-80 ms in order to avoid ocular exploration. When a target was detected (food or animal) the subject had to let go of the button in less than 1s otherwise he had to keep pressing the button for at least 1s. Stimuli were as varied as possible, targets were presented in their natural environment and included mammals, birds, fish, and reptiles in the animal task, fruits, vegetables, and cakes in the food task. Subject had no a priori information on the location of the target in the image, its size, the number of targets, or the view angle. Distractors included a large variety of natural scenes and man-made objects (landscapes, waterfalls, flowers, buildings and some targets of the other task...) so that all steps of the ventral visual pathway concerned by object recognition must be involved in solving the problem. Performance of two rhesus macaques and of two groups of 10 human subjects were studied on the same sets of stimuli. Humans were tested on series of 100 new stimuli whereas monkeys had to categorise daily 5 new targets and 5 new distractors that were randomly mixed in a sequence of familiar images that they had experienced several times before. However the only behavioural response considered in the analysis was the response given to the *first* presentation of a totally new image. Monkeys and humans were tested using two sets of 200 new images in both tasks. Macaques were slightly less accurate in their decision than humans but much faster. The monkeys scored 90.5% and 89.5% correct in the food task, 84% and 87% correct in the animal task, whereas human subjects

reached on average 95% in both cases. In contrast monkeys could produce their response with a median reaction time (RT) ranging from 233-344 ms whereas median RT in humans ranged in between 383-504 ms. These reaction times are very short even in humans and must include also a substantial motor component so that if categorisation involves all the processing steps along the ventral visual pathway from retina to inferotemporal cortex, then the visual processing must be massively parallel and essentially feed forward. Monkeys and humans would thus rely on the processing of the earliest visual information available. These kind of processing might be very similar in between the two species. In order to tackle this problem, we analysed all the errors that had been made by humans and monkeys. There is a considerable overlap of the two sets of errors. Monkeys categorise with a high accuracy images that were correctly classified by all humans, but fall around chance level for images that induced categorisation errors for more than 40% of the human subjects. Moreover, monkeys and humans performances are affected similarly (1) by training and (2) by the removal of colour cues. (1) Familiar images are classified more accurately but training cannot shorten the latency of the earliest responses. (2) When removing colour cues from the stimuli so that human and nonhuman primates have to categorise black and white images mixed randomly with colour images, monkeys performance is almost unimpaired in accuracy and in speed. This lack of effect is also seen in humans producing their motor response with short RT whereas humans with long RT appear to rely slightly more on colour cues. These results show that monkeys can perform extremely fast and reliable classifications of previously unseen stimuli (even without eye movements and without colour cues) probably on the basis of abstract representations that overlap with those of humans. Evolution could have set up in humans and monkeys a fast and very efficient way of processing information that could show large similarities in the two species.

BIBLIOGRAPHIE

- Aust, U., Loidolt, M., & Huber, L. (1999). Acquisition of a people concept versus use of simple features. *Perception, 28* (suppl), 21.
- Booth, M. C. A., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex, 8*, 510-523.
- Chadaide, Z., Kovács, G., Köteles, K., & Benedek, G. (1999). Selectivity of macaque inferior temporal neurons for shapes with different surface attributes. *Perception, 28* (suppl), 97.

- D'Amato, M. R., & Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 43-55.
- D'Amato, M. R., & Worsham, R. W. (1972). Delayed matching in the capuchin monkey with brief sample durations. *Learning and Motivation*, *3*, 304-312.
- Dasser, V. (1988). A social concept in Java Monkeys. *Animal Behaviour*, *36*, 225-230.
- Davenport, R. K., & Rogers, C. M. (1970). Intermodal equivalence of stimuli in apes. *Science*, *168*, 279-280.
- Delorme, A., Richard, G., & Fabre-Thorpe, M. (soumis). Ultra-rapid categorisation of natural images by humans and rhesus monkeys is color blind.
- Dulai, K. S., Bowmaker, J. K., Mollon, J. D., & Hunt, D. M. (1994). Sequence divergence, polymorphism and evolution of the middle-wave and long-wave visual pigment genes of great apes and old world monkeys. *Vision Research*, *34*, 2483-2491.
- Elliot, R. C. (1977). Cross-modal recognition in three primates. *Neuropsychologia*, *15*, 183-186.
- Fabre-Thorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *NeuroReport*, *9*, 303-308.
- Fischer, G. J., & Kitchener, S. L. (1962). Comparative learning in young gorillas and orang-utans. *Journal of Genetic Psychology*, *107*, 337-348.
- Frégnac, Y., & Bringuier, V. (1996). Spatio-temporal dynamics of synaptic integration in cat visual cortical receptive fields. In A. Aertsen & V. Braitenberg (Eds.), *Biological basis and computational theory of vision* (pp. 1-47). Amsterdam: Elsevier.
- Herrnstein, R. J. (1984). Objects, categories, and discriminative stimuli. In H. T. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 233-262). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, *146*, 549-551.
- Hopkins, W. D., Fagot, J., & Vauclair, J. (1993). Mirror-image matching and mental rotation problem solving in baboons (*Papio papio*): unilateral input enhances performance. *Journal of Experimental Psychology: General*, *122*, 61-72.
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552-563.
- Medin, D. L. (1969). Form perception and pattern reproduction by monkeys. *Journal of Comparative and Physiological Psychology*, *68*, 412-419.
- Milner, A. D., & Goodale, M. A. (1993). Visual pathways to perception and action. *Progress in Brain Research*, *95*, 317-337.
- Nowak, L. G., & Bullier, J. (1997). The timing of information transfer in the visual system. In J. Kaas, K. Rocklund, & A. Peters (Eds.), *Extrastriate cortex in primates* (pp. 205-241). New York: Plenum Press.
- Nowak, L. G., Munk, M. H. J., Girard, P., & Bullier, J. (1995). Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neuroscience*, *12*, 371-384.

- Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology*, *68*, 70-84.
- Owren, M. J. (1990a). Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): I. Natural calls. *Journal of Comparative Psychology*, *104*, 20-28.
- Owren, M. J. (1990b). Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): II. Synthetic calls. *Journal of Comparative Psychology*, *104*, 29-40.
- Parr, L.A., & de Waal, F.B.M. (1999). Visual kin recognition in chimpanzees. *Nature*, *399*, 647-648.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329-342.
- Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences*, *6*, 125-167.
- Roberts, W. A., & Mazmanian, D. S. (1988). Concept learning at different levels of abstraction by pigeons, monkeys, and people. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 247-260.
- Rumbaugh, D. M. (1970). Learning skills of anthropoids. Primate behavior: In L. A. Rosenblum (Ed.), *Developments in field and laboratory research* (pp. 1-70). New York: Academic Press.
- Rumbaugh, D. M., & Pate, J. L. (1984a). The evolution of cognition in primates: a comparative perspective. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 569-587). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rumbaugh, D. M., & Pate, J. L. (1984b). Primates' learning by levels. In G. Greenberg & E. Tobach (Eds.), *Behavioral evolution and integrative levels* (pp. 221-240). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schrier, A. M., Angarella, R., & Povar, M. L. (1984). Studies of concept formation by stumptailed monkeys: concepts humans, monkeys, and letter A. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 564-584.
- Schrier, A. M., & Brady, P. M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*): effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behavior Processes*, *13* (2), 136-143.
- Seyfarth, R. M. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, *210*, 801-803.
- Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*, 869-873.
- Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, *66*, 170-189.
- Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520-522.

- Thorpe, S. J., & Imbert, M. (1989). Biological constraints on connectionist models. In R. Pfeifer, Z. Schreter, F. Fogelman-Soulié, & L. Steels (Eds.), *Connectionism in perspective* (pp. 63-92). Amsterdam: Elsevier.
- Troje, N. F., Huber, L., Loidolt, M., Aust, U., & Fieder, M. (1999). Categorical learning in pigeons: the role of texture and shape in complex static stimuli. *Vision Research*, *39*, 353-366.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549-585). Cambridge, MA: MIT Press.
- Vauclair, J., Rollins Jr., H. A., & Nadler, R. D. (1983). Reproductive memory for diagonal and non-diagonal patterns in chimpanzees. *Behavioural Processes*, *8*, 289-300.
- Vogels, R. (1999). Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *European Journal of Neuroscience*, *11*, 1239-1255.
- Watanabe, S. (1997). Visual discrimination of real objects and pictures in pigeons. *Animal Learning Behavior*, *25*, 185-192.
- Weiskrantz, L., & Cowey, A. (1975). Cross-modal matching in the rhesus monkey using a single pair of stimuli. *Neuropsychologia*, *13*, 257-261.
- Yoshikubo, S. (1985). Species discrimination and concept formation by rhesus monkeys (*Macaca mulatta*). *Primates*, *26*, 285-299.

